A STOCHASTIC OPTIMAL CONTROL STRATEGY FOR PARTIALLY OBSERVABLE NONLINEAR SYSTEMS

W. Q. Zhu and Z. G. Ying

Department of Mechanics, Zhejiang University, Hangzhou 310027, People's Republic of China

Abstract: A stochastic optimal control strategy for partially observable nonlinear systems is proposed. The optimal control force consists of two parts. The first part is determined by the conditions under which the stochastic optimal control problem of a partially observable nonlinear system is converted into that of a completely observable linear system. The second part is determined by solving the dynamical programming equation derived by applying the stochastic averaging method and stochastic dynamical programming principle to the completely observable linear control system. For controlled quasi Hamiltonian systems, the response is predicted by solving the FPK equation and the Riccati equation.

Keywords: nonlinear system, partial observation, stochastic optimal control, separation principle, stochastic averaging, dynamical programming

1. INTRODUCTION

Stochastic optimal control is a research subject of much significance since many actual control systems such as those in engineering structures are subjected to random excitations and the system states are estimated from the measurements with random noises (Housner, *et al.*, 1997). For a long period of time, only the linear quadratic Gaussian (LQG) control strategy was used in engineering applications. In recent years, several optimal control strategies for stochastically excited nonlinear systems have been proposed (Bernstein, 1993; Crespo and Sun, 2003; Zhu, *et al.*, 2001). In these studies, the states of the controlled systems were assumed known exactly, i.e., the controlled systems are completely observable. However, the system states are actually estimated from the measurements with random noises, i.e., the controlled systems are partially observable. One basic approach to the stochastic optimal control of partially observable systems is to convert the stochastic optimal control problem of a partially observable system into that of a completely observable system using the separation principle (Wonham, 1968; Fleming and Rishel, 1975; Bensoussan, 1992) and then to solve the later problem. For a partially observable linear system, the converted completely observable control system is of finite dimension and it can be solved easily, e.g., by using LQG strategy. A nonlinear stochastic optimal control strategy for partially observable linear systems was proposed recently by present authors (Zhu and Ying, 2002) based on the separation principle, stochastic averaging method and stochastic dynamical programming principle. For a partially observable nonlinear system, the converted completely observable control system is usually of infinite dimension and it can hardly be solved. A few years ago, Charalambous and Elliott (1998) proved that if the nonlinearities enter the dynamics of the unobservable states and the observations as gradients of potential functions, then the partially observable nonlinear control system can be recast as a completely observable linear control system of finite dimension.

The objective of the present paper is to propose a nonlinear stochastic optimal control strategy for partially observable nonlinear systems based on the theorem due to Charalambous and Elliott (1998) and the nonlinear stochastic optimal control strategy proposed for completely observable systems by Zhu et al. (2001).

2. STOCHASTIC OPTIMAL CONTROL PROBLEM OF PARTIALLY OBSERVABLE NONLINEAR SYSTEMS

Consider a controlled nonlinear system governed by

$$\mathrm{d}\mathbf{X} = \overline{\mathbf{A}}(\mathbf{X})\mathrm{d}t + \overline{\mathbf{U}}(\mathbf{X})\mathrm{d}t + \mathbf{C}_1\mathrm{d}\mathbf{B}(t) \qquad (1)$$

where $\overline{\mathbf{A}}(\mathbf{X})$ and $\overline{\mathbf{U}}$ are $2n$-dimensional function vectors; $\mathbf{C}_1$ is $2n\times m$-dimensional matrix; $\mathbf{B}(t)$ is m-dimensional Wiener process vector. The observation equation is of the form

$$\mathrm{d}\mathbf{Y} = \overline{\mathbf{D}}(\mathbf{X})\mathrm{d}t + \mathbf{F}\overline{\mathbf{U}}\mathrm{d}t + \mathbf{C}_2\mathrm{d}\mathbf{B}(t) + \mathbf{C}_3\mathrm{d}\mathbf{B}_1(t) \qquad (2)$$

where $\mathbf{Y}$ is n_1-dimensional observation vector; $\mathbf{D}(\mathbf{X})$ is n_1-dimensional function vector; $\mathbf{B}_1(t)$ is m_1-dimensional Wiener process vector; $\mathbf{F}$, $\mathbf{C}_2$ and $\mathbf{C}_3$ are $n_1\times 2n$, $n_1\times m$ and $n_1\times m_1$-dimensional constant matrices, respectively. The objective of stochastic optimal control is to minimize a performance index

$$J = E\{\int_0^T L(\mathbf{X},\overline{\mathbf{U}})\mathrm{d}t + \Psi(\mathbf{X}(T))\} \qquad (3a)$$

for finite time-interval control, or

$$J' = \lim_{T\to\infty}\frac{1}{T}\int_0^T L(\mathbf{X},\overline{\mathbf{U}})\mathrm{d}t \qquad (3b)$$

for semi-infinite time-interval ergodic control, where $E\{\cdot\}$ denotes expectation operation; T is the terminal time of control; $L(\mathbf{X}, \mathbf{U})$ is cost function, which is continuous, differential and convex function; $\Psi(T)$ is terminal cost. Eqs. (1), (2) and (3) constitute a stochastic optimal control problem of partially observable nonlinear system. It consists of two coupled problems of optimal filtering and optimal control.

To convert this stochastic optimal control problem into one of completely observable linear system, control force $\overline{\mathbf{U}}$ is first split into $\overline{\mathbf{U}}_1$ and $\overline{\mathbf{U}}_2$. $\overline{\mathbf{U}}_1$ is combined with the uncontrolled system and observation so that Eqs. (1) and (2) become

$$\mathrm{d}\mathbf{X} = [\mathbf{A}\mathbf{X} + \mathbf{G}(\mathbf{X})]\mathrm{d}t + \overline{\mathbf{U}}_2\mathrm{d}t + \mathbf{C}_1\mathrm{d}\mathbf{B}(t) \qquad (4)$$

$$\mathrm{d}\mathbf{Y} = [\mathbf{D}\mathbf{X} + \mathbf{E}(\mathbf{X})]\mathrm{d}t + \mathbf{F}\overline{\mathbf{U}}_2\mathrm{d}t + \mathbf{C}_2\mathrm{d}\mathbf{B}(t) + \mathbf{C}_3\mathrm{d}\mathbf{B}_1(t) \qquad (5)$$

where $\mathbf{A}$ and $\mathbf{D}$ are $2n\times 2n$ and $n_1\times 2n$-dimensional constant matrices, respectively;

$$\mathbf{A} = \frac{\partial}{\partial\mathbf{X}}(\overline{\mathbf{A}}(0) + \overline{\mathbf{U}}_1(0)),\ \mathbf{G}(\mathbf{X}) = \overline{\mathbf{A}}(\mathbf{X}) + \overline{\mathbf{U}}_1 - \mathbf{A}\mathbf{X},$$

$$\mathbf{D} = \frac{\partial}{\partial\mathbf{X}}(\overline{\mathbf{D}}(0) + \mathbf{F}\overline{\mathbf{U}}_1(0)),\ \mathbf{E}(\mathbf{X}) = \overline{\mathbf{D}}(\mathbf{X}) + \mathbf{F}\overline{\mathbf{U}}_1 - \mathbf{D}\mathbf{X} \qquad (6)$$

Note that control system (4) and observation (5) contain nonlinear terms $\mathbf{G}(\mathbf{X})$ and $\mathbf{E}(\mathbf{X})$, respectively. Correspondingly, performance index (3) is modified as

$$J_0 = E\{\int_0^T L(\mathbf{X},\overline{\mathbf{U}}_2)\mathrm{d}t + \Psi(\mathbf{X}(T))\} \qquad (7a)$$

for finite time-interval control, or

$$J_0' = \lim_{T\to\infty}\frac{1}{T}\int_0^T L(\mathbf{X},\overline{\mathbf{U}}_2)\mathrm{d}t \qquad (7b)$$

for semi-infinite time-interval ergodic control.

3. CONVERTED STOCHASTIC OPTIMAL CONTROL PROBLEM OF COMPLETELY OBSERVABLE LINEAR SYSTEMS

According to the separation principle (Wonham, 1968; Fleming and Rishel, 1975; Bensoussan, 1992), the stochastic optimal control problem of partially observable system (4), (5) and (7) can be converted into one of completely observable system. However, usually it is of infinite dimension and thus a very difficult problem. To make the converted stochastic optimal control problem of completely observable system of finite dimension, according to Charalambous and Elliott (1998), assume that initial system state $\hat{\mathbf{X}}(0)$ has the following probability density

$$p_0(\hat{\mathbf{X}}|\mathbf{Y}) = \frac{\mathrm{e}^{-(\hat{\mathbf{X}}-\mathbf{m}_0)^T\boldsymbol{\sigma}_0^{-1}(\hat{\mathbf{X}}-\mathbf{m}_0)/2}}{\sqrt{(2\pi)^n|\boldsymbol{\sigma}_0|}}\times\mathrm{e}^{\phi(\hat{\mathbf{X}},0)} \qquad (8)$$

where $\mathbf{m}_0$ and $\boldsymbol{\sigma}_0$ are constant vector and symmetric positive-definite matrix, respectively, and $\overline{\mathbf{U}}_1$ is selected so that the nonlinear terms in control system

(4) and observation (5) have potential function $\phi(\hat{\mathbf{X}},t)$, i.e.,

$$\mathbf{G}(\hat{\mathbf{X}}) = \mathbf{C}_1\mathbf{C}_1^T \frac{\partial\phi(\hat{\mathbf{X}},t)}{\partial\hat{\mathbf{X}}}, \quad \mathbf{E}(\hat{\mathbf{X}}) = \mathbf{C}_2\mathbf{C}_1^T \frac{\partial\phi(\hat{\mathbf{X}},t)}{\partial\hat{\mathbf{X}}} \tag{9}$$

where $\phi(\hat{\mathbf{X}},t)$ satisfies the following partial differential equation

$$\begin{aligned}&\frac{\partial\phi}{\partial t} + \frac{1}{2}\mathrm{tr}(\mathbf{C}_1\mathbf{C}_1^T \frac{\partial^2\phi}{\partial\hat{\mathbf{X}}^2}) + \frac{1}{2}\left|\mathbf{C}_1^T \frac{\partial\phi}{\partial\hat{\mathbf{X}}}\right|^2 \\ &\quad + (\mathbf{A}\hat{\mathbf{X}} + \overline{\mathbf{U}}_2)^T \frac{\partial\phi}{\partial\hat{\mathbf{X}}} = 0\end{aligned} \tag{10}$$

$\hat{\mathbf{X}}$ is system state estimation for given observation $\mathbf{Y}(\tau)$, $0\le\tau\le t$. If theses conditions are satisfied, then the stochastic optimal control problem of partially observable system (4), (5) and (7) can be converted into a stochastic optimal control problem of completely observable linear system, and formulated as

$$\mathrm{d}\hat{\mathbf{X}} = (\mathbf{A}\hat{\mathbf{X}} + \overline{\mathbf{U}}_2)\mathrm{d}t + (\mathbf{R}_C\mathbf{D}^T + \mathbf{C}_1\mathbf{C}_2^T)\mathbf{C}^{-1}\mathrm{d}\mathbf{V}_I \tag{11}$$

$$\mathrm{d}\mathbf{V}_I = \mathrm{d}\mathbf{Y} - \mathbf{D}\hat{\mathbf{X}}\mathrm{d}t \tag{12}$$

$$J_2 = E\{\int_0^T L_2(\hat{\mathbf{X}}, \overline{\mathbf{U}}_2)\mathrm{d}t + \Psi_2(\hat{\mathbf{X}}(T))\} \tag{13a}$$

for finite time-interval control, or

$$J_2' = \lim_{T\to\infty}\frac{1}{T}\int_0^T L_2(\hat{\mathbf{X}}, \overline{\mathbf{U}}_2)\mathrm{d}t \tag{13b}$$

for semi-infinite time-interval ergodic control, where $\mathbf{C} = \mathbf{C}_2\mathbf{C}_2^T + \mathbf{C}_3\mathbf{C}_3^T$; $\mathbf{V}_I$ is n_1-dimensional innovation process vector; $\mathbf{R}_C$ is the covariance matrix of state estimation error $\tilde{\mathbf{X}} = \mathbf{X} - \hat{\mathbf{X}}$, which has Gaussian probability density

$$p(\tilde{\mathbf{X}}) = \frac{1}{\sqrt{(2\pi)^n|\mathbf{R}_C|}}\mathrm{e}^{-\tilde{\mathbf{X}}^T\mathbf{R}_C^{-1}\tilde{\mathbf{X}}/2} \tag{14}$$

Covariance $\mathbf{R}_C$ satisfies the following differential Riccati equation

$$\begin{aligned}\dot{\mathbf{R}}_C &= \mathbf{A}\mathbf{R}_C + \mathbf{R}_C\mathbf{A}^T - (\mathbf{R}_C\mathbf{D}^T + \mathbf{C}_1\mathbf{C}_2^T) \\ &\quad \cdot\mathbf{C}^{-1}(\mathbf{D}\mathbf{R}_C + \mathbf{C}_2\mathbf{C}_1^T) + \mathbf{C}_1\mathbf{C}_1^T\end{aligned} \tag{15a}$$

for finite time-interval control, or algebraic Riccati equation

$$\begin{aligned}&\mathbf{A}\mathbf{R}_C + \mathbf{R}_C\mathbf{A}^T - (\mathbf{R}_C\mathbf{D}^T + \mathbf{C}_1\mathbf{C}_2^T) \\ &\quad \cdot\mathbf{C}^{-1}(\mathbf{D}\mathbf{R}_C + \mathbf{C}_2\mathbf{C}_1^T) + \mathbf{C}_1\mathbf{C}_1^T = 0\end{aligned} \tag{15b}$$

for semi-infinite time-interval ergodic control.

4. OPTIMAL CONTROL LAW

The stochastic optimal control problem of completely observable linear system (11)-(13) can be solved by using LQG control strategy. However, it has been shown that the nonlinear stochastic optimal control strategy proposed by the present authors (Zhu, *et al.*, 2001) is superior than LQG controller, especially, more effective and efficient. To apply the nonlinear stochastic optimal control strategy, let $\mathbf{X}=[\mathbf{Q}^T, \mathbf{P}^T]^T$ and system (1) is formulated as a controlled quasi Hamiltonian system, i.e.,

$$\overline{\mathbf{A}}(\mathbf{X}) = \begin{bmatrix}\partial H/\partial\mathbf{P} \\ -\partial H/\partial\mathbf{Q} - \mathbf{C}_0\partial H/\partial\mathbf{P}\end{bmatrix},$$

$$\overline{\mathbf{U}} = \begin{bmatrix}0 \\ \mathbf{U}\end{bmatrix}, \quad \mathbf{C}_1 = \begin{bmatrix}0 \\ \mathbf{K}_0\end{bmatrix} \tag{16}$$

where $\mathbf{Q}$ and $\mathbf{P}$ are n-dimensional generalized displacement and momentum vectors, respectively; $H=H(\mathbf{Q},\mathbf{P})$ is Hamiltonian, possibly modified by Wong-Zakai correction terms; $\mathbf{U}=\mathbf{U}(\mathbf{Q},\mathbf{P})$ is n-dimensional feedback control force vector; $\mathbf{C}_0=\mathbf{C}_0(\mathbf{Q},\mathbf{P})$ is $n\times n$-dimensional damping coefficient matrix possibly modified by Wong-Zakai correction terms; $\mathbf{K}_0=\mathbf{K}_0(\mathbf{Q},\mathbf{P})$ is $n\times m$-dimensional stochastic excitation amplitude matrix; $\mathbf{C}_0$, $\mathbf{U}$ and $\mathbf{K}_0\mathbf{K}_0^T$ are assumed of the same small order. In this case,

$$\mathbf{A} = \begin{bmatrix}\dfrac{\partial^2\overline{H}(0)}{\partial\mathbf{Q}\partial\mathbf{P}} & \dfrac{\partial^2\overline{H}(0)}{\partial\mathbf{P}^2} \\ -\dfrac{\partial^2\overline{H}(0)}{\partial\mathbf{Q}^2} - \dfrac{\partial}{\partial\mathbf{Q}} & -\dfrac{\partial^2\overline{H}(0)}{\partial\mathbf{P}\partial\mathbf{Q}} - \dfrac{\partial}{\partial\mathbf{P}} \\ (\mathbf{C}_0(0)\dfrac{\partial\overline{H}(0)}{\partial\mathbf{P}}) & (\mathbf{C}_0(0)\dfrac{\partial\overline{H}(0)}{\partial\mathbf{P}})\end{bmatrix} \tag{17}$$

where $\overline{H}$ is the Hamiltonian modified by $\mathbf{U}_1$. Eqs. (9) and (10) become

$$\begin{aligned}&(\frac{\partial\overline{H}}{\partial\hat{\mathbf{Q}}} + \mathbf{C}_0\frac{\partial\overline{H}}{\partial\hat{\mathbf{P}}})_N = -\mathbf{K}_0\mathbf{K}_0^T\frac{\partial\phi}{\partial\hat{\mathbf{P}}}, \\ &(\overline{\mathbf{D}}(\hat{\mathbf{X}}) + \mathbf{F}\overline{\mathbf{U}}_1)_N = \mathbf{C}_2\mathbf{K}_0^T\frac{\partial\phi}{\partial\hat{\mathbf{P}}},\end{aligned} \tag{18}$$

$$\begin{aligned}&\frac{\partial\phi}{\partial t} + \frac{1}{2}\mathrm{tr}(\mathbf{K}_0\mathbf{K}_0^T\frac{\partial^2\phi}{\partial\hat{\mathbf{P}}^2}) + \frac{1}{2}\left|\mathbf{K}_0^T\frac{\partial\phi}{\partial\hat{\mathbf{P}}}\right|^2 + (\frac{\partial\overline{H}}{\partial\hat{\mathbf{P}}})^T \\ &\cdot\frac{\partial\phi}{\partial\hat{\mathbf{Q}}} - (\frac{\partial\overline{H}}{\partial\hat{\mathbf{Q}}} + \mathbf{C}_0\frac{\partial\overline{H}}{\partial\hat{\mathbf{P}}})_L^T\frac{\partial\phi}{\partial\hat{\mathbf{P}}} + \mathbf{U}_2^T\frac{\partial\phi}{\partial\hat{\mathbf{P}}} = 0\end{aligned} \tag{19}$$

where $(\cdot)_N$ and $(\cdot)_L$ represent nonlinear and linear terms, respectively. For stationary potential $\phi(\hat{\mathbf{X}})$, the first term in Eq. (19) vanishes. By applying the stochastic averaging method (Zhu, *et al.*, 1997) to system (11), the following averaged Itô stochastic differential equations is obtained

$$\mathrm{d}\hat{\mathbf{H}} = [\mathbf{m}(\hat{\mathbf{H}}) + < (\frac{\partial \hat{\mathbf{H}}}{\partial \hat{\mathbf{P}}})^T \mathbf{U}_2 >]\mathrm{d}t + \boldsymbol{\sigma}(\hat{\mathbf{H}})\mathrm{d}\mathbf{B}_3(t) \quad (20)$$

where $\hat{\mathbf{H}} = [\hat{H}_1, \hat{H}_2, \cdots, \hat{H}_n]^T$ and $\hat{H}_i$ is the ith modal energy of the controlled linear system; <·> denotes averaging operation; $\mathbf{B}_3(t)$ is standard Wiener process vector; $\mathbf{m}(\hat{\mathbf{H}})$ and $\boldsymbol{\sigma}(\hat{\mathbf{H}})$ are, respectively, drift vector and diffusion matrix with elements

$$\begin{aligned} m_i(\hat{\mathbf{H}}) =& < -\sum_{j,k=1}^{n} \bar{c}_{jk} \frac{\partial \hat{H}_i}{\partial \hat{P}_j} \frac{\partial \hat{H}}{\partial \hat{P}_k} + \int_{-\infty}^{0} \sum_{k,l=1}^{n_1} \sum_{j=1}^{n} \sum_{r,s=1}^{2n} \\ & [(\frac{\partial \hat{H}_j}{\partial \hat{X}_s} f_{sl})_{t+\tau} \frac{\partial}{\partial \hat{H}_j} (\frac{\partial \hat{H}_i}{\partial \hat{X}_r} f_{rk})_t \\ & + (\frac{\partial \hat{\theta}_j}{\partial \hat{X}_s} f_{sl})_{t+\tau} \frac{\partial}{\partial \hat{\theta}_j} (\frac{\partial \hat{H}_i}{\partial \hat{X}_r} f_{rk})_t] R_{kl}(\tau)\mathrm{d}\tau > \end{aligned} \quad (21)$$

$$\begin{aligned} \sigma_{i\cdot}(\hat{\mathbf{H}})\sigma_{j\cdot}(\hat{\mathbf{H}}) =& < \int_{-\infty}^{\infty} \sum_{k,l=1}^{n_1} \sum_{r,s}^{2n} (\frac{\partial \hat{H}_j}{\partial \hat{X}_s} f_{sl})_{t+\tau} \\ & (\frac{\partial \hat{H}_i}{\partial \hat{X}_r} f_{rk})_t R_{kl}(\tau)\mathrm{d}\tau > \end{aligned} \quad (22)$$

in which $\bar{c}_{jk}$ is damping coefficient dependent on **A** in Eq. (17); $\hat{\theta}_j$ is generalized phase process; f_{rk} is the element of matrix $(\mathbf{R}_C \mathbf{D}^T + \mathbf{C}_1 \mathbf{C}_2^T)\mathbf{C}^{-1}$; $R_{kl}(\tau)$ is the correlation function of $\mathbf{V}_l(t)$. Eq. (20) implies that $\hat{\mathbf{H}}(t)$ is a controlled diffusion process vector. Correspondingly, performance index (13) is modified into

$$J_3 = E\{\int_0^T < L_3(\hat{\mathbf{H}}, \mathbf{U}_2) > \mathrm{d}t + \Psi_3(\hat{\mathbf{H}}(T))\} \quad (23)$$

for finite time-interval control, or

$$J_4 = \lim_{T\to\infty} \frac{1}{T} \int_0^T < L_3(\hat{\mathbf{H}}, \mathbf{U}_2) > \mathrm{d}t \quad (24)$$

for semi-infinite time-interval ergodic control.

By applying the stochastic dynamical programming principle (Fleming and Rishel, 1975; Fleming and Soner, 1992) to the control problem of averaged system (20) and (23) or (24), a dynamical programming equation can be established. For performance index (23), it is

$$\begin{aligned} \frac{\partial V_1}{\partial t} =& -\min_{\mathbf{U}_2} \{< L_3(\hat{\mathbf{H}}, \mathbf{U}_2) > + [\mathbf{m}(\hat{\mathbf{H}}) + < (\frac{\partial \hat{\mathbf{H}}}{\partial \hat{\mathbf{P}}})^T \\ & \cdot \mathbf{U}_2 >]^T \frac{\partial V_1}{\partial \hat{\mathbf{H}}} + \frac{1}{2} \mathrm{tr}(\boldsymbol{\sigma}(\hat{\mathbf{H}})\boldsymbol{\sigma}^T(\hat{\mathbf{H}}) \frac{\partial^2 V_1}{\partial \hat{\mathbf{H}}^2})\} \end{aligned} \quad (25)$$

and for performance index (24), it is

$$\begin{aligned} \lambda =& \min_{\mathbf{U}_2} \{< L_3(\hat{\mathbf{H}}, \mathbf{U}_2) > + [\mathbf{m}(\hat{\mathbf{H}}) + < (\frac{\partial \hat{\mathbf{H}}}{\partial \hat{\mathbf{P}}})^T \\ & \cdot \mathbf{U}_2 >]^T \frac{\partial V_2}{\partial \hat{\mathbf{H}}} + \frac{1}{2} \mathrm{tr}(\boldsymbol{\sigma}(\hat{\mathbf{H}})\boldsymbol{\sigma}^T(\hat{\mathbf{H}}) \frac{\partial^2 V_2}{\partial \hat{\mathbf{H}}^2})\} \end{aligned} \quad (26)$$

In Eqs. (25) and (26), V_1 and V_2 are called value function, and λ is a constant representing optimal average cost. The optimal control law is determined by minimizing the right-hand side of Eq. (25) or (26) with respect to $\mathbf{U}_2$, i.e.,

$$\frac{\partial L_3}{\partial \mathbf{U}_2} + \frac{\partial \hat{\mathbf{H}}}{\partial \hat{\mathbf{P}}} \frac{\partial V}{\partial \hat{\mathbf{H}}} = 0 \quad (27)$$

Let L_3 be quadratic with respect to $\mathbf{U}_2$, i.e.,

$$L_3(\hat{\mathbf{H}}, \mathbf{U}_2) = g(\hat{\mathbf{H}}) + \mathbf{U}_2^T \mathbf{R} \mathbf{U}_2 \quad (28)$$

where $g(\hat{\mathbf{H}}) \geq 0$; **R** is a symmetric positive-definite matrix. Then optimal control law is of the form

$$\mathbf{U}_2^* = -\frac{1}{2} \mathbf{R}^{-1} \frac{\partial \hat{\mathbf{H}}}{\partial \hat{\mathbf{P}}} \frac{\partial V}{\partial \hat{\mathbf{H}}} \quad (29)$$

which depends on the derivatives of value function V_1 or V_2 with respect to $\hat{\mathbf{H}}$. Substituting Eq. (29) into Eq. (25) or (26) yields the following final dynamical programming equation

$$\begin{aligned} & \frac{\partial V}{\partial t} + \frac{1}{2} \mathrm{tr}(\boldsymbol{\sigma}\boldsymbol{\sigma}^T \frac{\partial^2 V}{\partial \hat{\mathbf{H}}^2}) + \mathbf{m}^T \frac{\partial V}{\partial \hat{\mathbf{H}}} - \frac{1}{4} \\ & \cdot < (\frac{\partial \hat{\mathbf{H}}}{\partial \hat{\mathbf{P}}} \frac{\partial V}{\partial \hat{\mathbf{H}}})^T R^{-1} \frac{\partial \hat{\mathbf{H}}}{\partial \hat{\mathbf{P}}} \frac{\partial V}{\partial \hat{\mathbf{H}}} > + g(\hat{\mathbf{H}}) = 0 \end{aligned} \quad (30)$$

in the case of finite time-interval control, or

$$\begin{aligned} & \frac{1}{2} \mathrm{tr}(\boldsymbol{\sigma}\boldsymbol{\sigma}^T \frac{\partial^2 V}{\partial \hat{\mathbf{H}}^2}) + \mathbf{m}^T \frac{\partial V}{\partial \hat{\mathbf{H}}} - \frac{1}{4} < (\frac{\partial \hat{\mathbf{H}}}{\partial \hat{\mathbf{P}}} \\ & \cdot \frac{\partial V}{\partial \hat{\mathbf{H}}})^T R^{-1} \frac{\partial \hat{\mathbf{H}}}{\partial \hat{\mathbf{P}}} \frac{\partial V}{\partial \hat{\mathbf{H}}} > + g(\hat{\mathbf{H}}) = \lambda \end{aligned} \quad (31)$$

in the case of semi-infinite time-interval ergodic control. Since the diffusion matrix in Eq. (30) or (31) is non-singular, they have classical solution (Fleming and Soner, 1992), i.e., continuous and smooth solution, which can be obtained by using conventional numerical technique. Thus, the second part of stochastic optimal control force, $\mathbf{U}_2^*$, can be obtained by solving Eq. (30) or (31) and then by substituting the resultant $\partial V / \partial \hat{\mathbf{H}}$ into Eq. (29). The total optimal control force is then $\mathbf{U}^* = \mathbf{U}_1 + \mathbf{U}_2^*$.

5. PERFORMANCE OF PROPOSED CONTROL STRATEGY

To evaluate the performance of the proposed stochastic optimal control strategy for partially observable nonlinear quasi Hamiltonian systems, the response of the optimally controlled system is first predicted. The response consists of optimally controlled response estimation $\hat{\mathbf{X}}$ and response estimation error $\tilde{\mathbf{X}}$. Substituting $\mathbf{U}_2^*$ in Eq. (29) into Eq. (20) and averaging the terms involving $\mathbf{U}_2^*$ yield

$$\mathrm{d}\hat{\mathbf{H}} = \overline{\mathbf{m}}(\hat{\mathbf{H}})\mathrm{d}t + \boldsymbol{\sigma}(\hat{\mathbf{H}})\mathrm{d}\mathbf{B}_3(t) \qquad (32)$$

where

$$\overline{\mathbf{m}}(\hat{\mathbf{H}}) = \mathbf{m}(\hat{\mathbf{H}}) + < (\frac{\partial \hat{\mathbf{H}}}{\partial \hat{\mathbf{P}}})^T \mathbf{U}_2^* > \qquad (33)$$

Solving the FPK equation associated with Itô equation (32) yields probability density $p(\hat{\mathbf{H}}, t)$ and mean square values $E[\hat{Q}_i^2]$ and $E[\hat{P}_i^2]$. Solving Riccati equation (15) yields the mean square values of the errors of estimated generalized displacements $E[\tilde{Q}_i^2] = (\mathbf{R}_C)_{ii}$ and momenta $E[\tilde{P}_i^2] = (\mathbf{R}_C)_{n+i,n+i}$. The total mean square generalized displacements and momenta are

$$E[Q_i^2] = E[\hat{Q}_i^2] + E[\tilde{Q}_i^2], \quad E[P_i^2] = E[\hat{P}_i^2] + E[\tilde{P}_i^2] \qquad (34)$$

from which the mean Hamiltonian $E[H_C]$ of the optimally controlled system can be obtained. The mean Hamiltonian $E[H_{UC}]$ of the uncontrolled system can be obtained by directly applying the stochastic averaging method to Eq. (1) with $\overline{\mathbf{U}}$ =0. The control effectiveness is measured in terms of

$$K_1 = \frac{E[H_{UC}] - E[H_C]}{E[H_{UC}]} \times 100\% \qquad (35)$$

$$K_2 = \frac{E[H_{UF}] - E[H_C]}{E[H_{UF}]} \times 100\% \qquad (36)$$

where $E[H_{UF}]$ is the mean Hamiltonian $E[H_F]$ for completely observable systems (1) and (3) plus the contribution from measurement error. Higher values of K_1 and K_2 imply better effectiveness of the proposed control strategy.

6. EXAMPLE

The proposed stochastic optimal control strategy has been applied to the ergodic control of a partially observable Duffing oscillator subjected to Gaussian white noise excitation governed by system equation

$$\ddot{X}_1 + c\dot{X}_1 + aX_1 + bX_1^3 = e\xi(t) + u \qquad (37)$$

where c=0.1, b/a=0.16, and observation equation

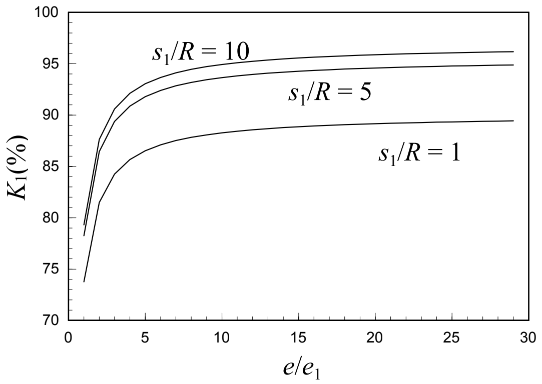

Fig. 1. K_1 as function of e/e_1 with s_1/R as parameter.

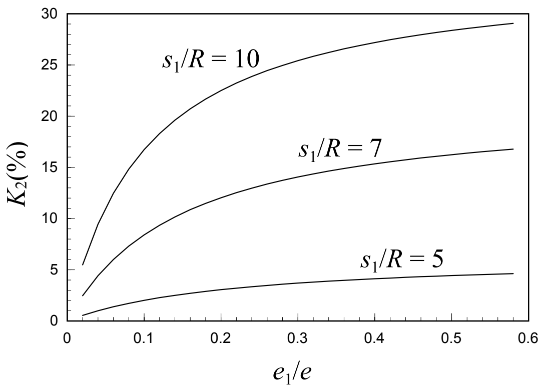

Fig. 2. K_2 as function of e_1/e with s_1/R as parameter.

$$\dot{Y} = \dot{X}_1 + e_1\xi_1(t) \qquad (38)$$

Some numerical results are shown in Figs. 1-2, which illustrate that the proposed control strategy is very effective even for large observation noise.

7. CONCLUSIONS

In the present paper, a stochastic optimal control strategy for partially observable nonlinear systems has been proposed. It has been shown through applying the proposed control strategy to a partially observable Duffing oscillator under stochastic excitation that the proposed control strategy is very effective even for large observation noise.

ACKNOWLEDGEMENTS

This study was supported by the National Natural Science Foundation of China under key grant No. 10332030, the Zhejiang Provincial Natural Science Foundation under grant No. 101046 and the Special Fund for Doctor Programs in Institutions of Higher Learning of China under grant No. 20020335092.

REFERENCES

Bensoussan, A. (1992). *Stochastic Control of Partially Observable Systems*, Cambridge University Press, Cambridge.

Bernstein, D.S. (1993). Nonquadratic cost and nonlinear feedback control. *Int. J. Robust Nonlinear Control*, 3, pp. 211-229.

Charalambous, C.D. and R.J. Elliott (1998). Classes of nonlinear partially observable stochastic optimal control problems with explicit optimal control laws. *SIAM J. Control Optim.*, 36, pp. 542-578.

Crespo, L.G. and J.Q. Sun (203). Nonlinear stochastic control via stationary response design. *Probab. Eng. Mech.*, 18, pp. 79-86.

Fleming, W.H. and H.M., Soner (1992). *Controlled Markov Processes and Viscosity Solutions*, Springer-Verlag, New York.

Fleming, W.H. and R.W. Rishel (1975). *Deterministic and Stochastic Optimal Control*, Springer-Verlag, Berlin.

Housner, G.W., L.A. Bergman, T.K. Caughey, A.G. Chassiakos, R.O. Claus, S.F. Masri, R.E. Skelton, T.T. Soong, B.F. Spencer and J.T.P. Yao (1997). Structural control: past, present, and future. *ASCE J. Eng. Mech.*, 123, pp. 897-971.

Stengel, R.F. (1986). *Stochastic Optimal Control: Theory and Application*, John Wiley & Sons, New York.

Wonham, W.M. (1968). On the separation theorem of stochastic control. *SIAM J. Control*, 6, pp. 312-326.

Zhu, W.Q., Z.L. Huang and Y.Q. Yang (1997). Stochastic averaging of quasi-integrable Hamiltonian systems. *ASME J. Appl. Mech.*, 64, pp. 975-984.

Zhu, W.Q. and Z.G. Ying (2002). Nonlinear stochastic optimal control of partially observable linear structures. *Eng. Struct.*, 24, pp. 333-342.

Zhu, W.Q., Z.G. Ying and T.T. Soong (2001). An optimal nonlinear feedback control strategy for randomly excited structural systems. *Nonlinear Dyn.*, 24, pp. 31-51.